



DATA PAPER

An updated global atmospheric paleo-reanalysis covering the last 400 years

Veronika Valler^{1,2} | Jörg Franke^{1,2} | Yuri Brugnara^{1,2} | Stefan Brönnimann^{1,2}

¹Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland

²Institute of Geography, University of Bern, Bern, Switzerland

Correspondence

Veronika Valler, Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland.
Email: veronika.valler@giub.unibe.ch

Funding information

H2020 European Research Council, Grant/Award Number: 787574; Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 162668

Abstract

Data assimilation techniques are becoming increasingly popular for climate reconstruction. They benefit from estimating past climate states from both observation information and from model simulations. The first monthly global paleo-reanalysis (EKF400) was generated over the 1600 and 2005 time period, and it provides estimates of several atmospheric fields. Here we present a new, considerably improved version of EKF400 (EKF400v2). EKF400v2 uses atmospheric-only general circulation model simulations with a greatly extended observational network of early instrumental temperature and pressure data, documentary evidences and tree-ring width and density proxy records. Furthermore, new observation types such as monthly precipitation amounts, number of wet days and coral proxy records were also included in the assimilation. In the version 2 system, the assimilation process has undergone methodological improvements such as the background-error covariance matrix is estimated with a blending technique of a time-dependent and a climatological covariance matrices. In general, the applied modifications resulted in enhanced reconstruction skill compared to version 1, especially in precipitation, sea-level pressure and other variables beside the mostly assimilated temperature data, which already had high quality in the previous version. Additionally, two case studies are presented to demonstrate the applicability of EKF400v2 to analyse past climate variations and extreme events, as well as to investigate large-scale climate dynamics.

KEYWORDS

climate reconstruction, ensemble Kalman fitting, paleoclimate data assimilation

Dataset

Identifier: 10.26050/WDCC/EKF400_v2.0

Creator: Jörg Franke, Veronika Valler, Yuri Brugnara, Stefan Brönnimann

Title: Ensemble Kalman Fitting Paleo-Reanalysis Version 2 (EKF400_v2)

Publisher: World Data Center for Climate (WDCC) at DKRZ

Publication year: 2020

Resource type: Dataset

Version: 2.0

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Geoscience Data Journal* published by Royal Meteorological Society and John Wiley & Sons Ltd.

1 | INTRODUCTION

Climate variability can be examined over several temporal and spatial timescales. To study climate variation at decadal to centennial timescales, long time series are required. Previous climate reconstructions suggest that strong natural variability characterized the period before 1850, but the underlying mechanisms are not fully understood (e.g. Brönnimann et al., 2019b). Before 1850, instrumental data are only available at limited locations and proxy records typically provide information at annual or lower resolution. Therefore, in many studies the different data types are combined to achieve a better performance. Multiple European reconstructions, for instance, combined instrumental measurements, historical evidence and proxy records (Luterbacher et al., 2004; Pauling et al., 2006; Küttel et al., 2010). These reconstructions with high spatial and temporal resolution can reveal regional differences and can identify distinct seasonal variability. However, such reconstructions are not available for the entire globe. Although not following the true evolution of past climate variability, model simulations provide information on forced and internal variability of various elements of the climate system at full spatial coverage. Therefore, they can contribute to better understand climate dynamics and related processes.

In addition to traditional reconstruction methods and climate models, data assimilation offers a new way to investigate past climate changes. In the paleoclimate data assimilation framework, past climate states are estimated from climate model simulations and observational information. The two sources of information are optimally combined, and the obtained reconstruction is regarded as the best estimate of the past climate. In the recent years, several studies explored different data assimilation methods to reconstruct past climate (e.g. Hakim et al., 2013, and references therein). For the climate of the last Millennium, the most used approaches rely on ensemble-based Kalman filtering techniques, which were applied, for instance, in pseudo-proxy experiments (e.g. Bhend et al., 2012), hydroclimate reconstructions (Steiger et al., 2018) or the generation of a monthly atmospheric paleo-reanalysis (EKF400, Franke et al., 2017).

EKF400 is a global 3-dimensional climate reconstruction, that spans the period between 1600 and 2005 (Franke et al., 2017). It was generated with the ensemble Kalman fitting technique (Franke et al., 2017), the offline variant of the ensemble square root filter (Whitaker & Hamill, 2002). It blends the ECHAM5.4 atmospheric model simulations called CCC400 (Bhend et al., 2012) with different types of observations (early instrumental measurements, documentary data, proxy records). EKF400 is the first paleo-reanalysis that provides multivariate information about the state of the atmosphere during the last 400 years at monthly resolution. Such data are highly relevant for climate impact studies or

to investigate the role of atmospheric circulation. It has been used, for instance, to study climate anomalies related to volcanic eruptions (Brönnimann et al., 2019b; Zambri, et al., 2019) and droughts (Burgdorf et al., 2019).

Given the novelty of EKF400, several further development possibilities were identified. This paper outlines the developments applied in the generation of the new version of EKF400 (EKF400v2). The observational network of previously assimilated data types is notably extended. In addition, in version 2 of EKF400, new data sources such as corals and precipitation information (amounts and number of wet days) were assimilated (Valler et al., 2020). Another line of improvements focuses on methodological developments. The new version of EKF400 takes advantages of improved covariance localization and the application of better estimated background-error covariance matrix (Valler et al., 2019).

The aim of this paper is to describe the dataset, highlight the improvements over EKF400 version 1 (EKF400v1) and to inform the users about its limitations, so that it can be effectively used in future climate studies. This article is organized as follows. In Section 2, we describe the model simulations and the observational network used as well as the ensemble Kalman fitting (EKF) paleoclimate data assimilation method. In Section 3, the reconstruction experiment and the methodological improvements are introduced. The skill of EKF400v2 is analysed in the period between 1902 and 2002 in Section 4. In addition, two case studies of Central European droughts from 1726 to 1728, and the effect of La Niña in the 19th century is presented in Section 4. Finally, we draw our concluding remarks in Section 5.

2 | DATA PRODUCTION METHODS

2.1 | Data sources

2.1.1 | Model simulations

EKF400v2 uses the same model simulations (CCC400) as EKF400v1. The 30 ensemble members were generated with the ECHAM5.4 general circulation model (Roeckner et al., 2003). The model simulations cover the period between 1600 and 2005. Reconstructed solar irradiance, sea surface temperatures and land surface parameters, volcanic activity, concentrations of greenhouse gases and sulphate were used to run the model at T63 resolution with 31 vertical levels (described in more detail in Bhend et al., 2012). The forcings and boundary conditions used to produce the model simulations are summarized in Table 1. Some errors were detected in the representation of land surface classes, which mainly affected the temperature field over the land areas in the extratropical Northern Hemisphere, and unrealistic wind speeds were found at southern high latitudes (Franke et al., 2017;

TABLE 1 Summary of boundary conditions and forcings used for the CCC400 model simulations

Feature	References
sea surface temperature	Mann et al., (2009)
Niño-3.4 index	Cook et al., (2008)
climatological sea-ice	Rayner et al., (2003)
solar irradiance	Lean (2000)
land-use	Pongratz et al., (2008)
volcanic forcing	Crowley et al., (2008)
greenhouse gases	as used in Yoshimori et al., (2010) and references therein
sulphate	Koch et al., (1999)

Rohrer et al., 2018). Hence, we recommend not to use the dataset south of 60°S. The 6-hourly model output files were converted to monthly means. Several fields are updated by the assimilation: temperature at 2 m, precipitation, number of wet days, sea-level pressure, horizontal wind components at 850, 500 and 200 hPa, vertical motion at 500 hPa, geopotential height at 1,000, 500, 300, 200 and 100 hPa, blockings and cyclone frequency. In version 2, we provide a dataset with 2° resolution globally, double the resolution of version 1.

2.1.2 | Observational dataset

Various data types from proxy records to documentary data to instrumental measurements are involved in the assimilation.

In EKF400v1, only a small number of proxy records of tree-ring width (TRW) and maximum latewood density (MXD) were used (Franke et al., 2017). In EKF400v2, additional tree-ring proxy records were included from different databases such as the vast collection of tree-ring width chronologies of Breitenmoser et al., (2014), the N-TREND (Wilson et al., 2016) and the PAGES2k-2017 (Emile-Geay et al., 2017) data collections. In addition, selected tree-ring chronologies from the collection by Neukom et al., (2014) were also added to the input data (Table 2).

In version 2 of EKF400, assimilated proxy types were further extended with the assimilation of corals. Corals have annual to sub-annual resolution, which is ideal for our monthly resolved reconstruction and they cover low-latitude regions, where trees cannot serve as temperature proxies. Sub-annually resolved coral chronologies from the PAGES2k-2017 database are used in EKF400v2. The proxy input data file has 3,513 entries in total, which means more than 20 times more proxy data compared to EKF400v1 (Figure 1).

Temperature information from documentary data was also assimilated in EKF400v1. The five documentary data from Central Europe provide monthly information about the climate including the winter (or non-growing) season, an

advantage compared to proxy records. In the input file of EKF400v2, two further documentary data series from Europe (Fernandez-Fernandez et al., 2014; Loader et al., 2011) and five new data series from Japan (Aono, 2015; Aono & Kazui, 2008; Aono & Tani, 2014; Mikami, 1996; Ohba et al., 2013) are included, leading to a total number of 12 documentary data series.

In the EKF400v1, most of the temperature time series were obtained from the Global Historical Climatology Network-Monthly version 3 dataset (GHCN-Monthly v3, Lawrimore et al., 2011), with the earliest series starting in 1801. Because the main focus of the paleo-reanalysis is the pre-20th century period, only series starting before 1,880 were kept. In EKF400v2, GHCN-Monthly v3 is replaced with the merged data collection of the international surface temperature initiative (ISTI, Rennie et al., 2014). Like in the case of GHCN-Monthly v3 data, time series that start before 1880 were selected, leaving 619 records. The earliest series in the ISTI database starts in 1701. The additional time series, which were included in EKF400v1 from the Historical Instrumental Climatological Surface Time series of the Greater Alpine Region database (HISTALP, Auer et al., 2007; Bohm et al., 2010) and separately collected series from various sources (Franke et al., 2017), were also included in EKF400v2, together with eight further stations from Brazil (Dove, 1847), Canada (Slonosky, 2014), France (Rousseau, 2009; Rousseau, 2013), India (Dove, 1847), Italy (Dove, 1847), Russia (Dove, 1847), Spain (Barriendos et al., 1997) and the United States (Blodgett, 1857) (Table 2). Altogether, the temperature input file contains 701 time series.

The GHCN-Monthly dataset version 2 contains precipitation data of several thousands stations (Lawrimore et al., 2011), from which the stations where measurements started before 1880 were included in the assimilation (1,437 stations). From the HISTALP data collection, 69 precipitation time series were used. In addition, 9 time series from Brazil (Farrona et al., 2012), England (Derham, 1705; Townley, 1699; Townley, 1694), France (Slonosky, 2002), Gibraltar (Wheeler, 2007), Ireland (Murphy et al., 2018), Jamaica, (Chenoweth, 2003), Madeira (Alcoforado et al., 2012), Spain (Barriendos, et al., 2002) and the United States (Blodgett, 1857) were added to the assimilated data (Table 2). In total, the precipitation input file contains 1515 stations data. In addition to monthly precipitation amounts, the number of wet days was also included in the assimilation. Wet days records were calculated from the GHCN-Daily dataset version 3.27 (Menne et al., 2012). A daily precipitation amount ≥ 1 mm was considered a wet day. A total of 652 wet days series are included, the earliest starting in 1781.

In the input file of EKF400v1, pressure time series were obtained from GHCN-Monthly v2 (Peterson & Vose, 1997), from HISTALP, from the collection made by Kuttel et al., (2010) and

TABLE 2 Input data used for EKF400v1 and EKF400v2

Data	Variable	Time	Data source	Reference
Anjarakandy*	T	1810–1823	inst	Dove (1847)
Archangelsk	T	1813–2001	inst	Brohan et al., (2006)
Barcelona*	T	1780–2004	inst	Barriendos et al., (1997)
Burnley*	R	1677–1704	inst	Townley, 1694, Townley, 1699), Derham (1705)
Cambridge* (US)	P	1742–1812	inst	Dupigny-Giroux et al., (2007)
Carpathian basin	T	1601–1854	docu	Bartholy et al., (2004)
Central Belgium	T	1794–2001	inst	Demarée et al., (2002)
Central England	T	1659–2004	inst	Manley (1974), Parker et al., (1992)
Central Europe	T	1601–1854	docu, inst	Dobrovolný et al., (2010)
Central Europe (Germany)	T	1600–1759	docu	Glaser and Riemann (2009)
Charleston*	T, R	1738–1759	inst	Blodget (1857)
Edo (Tokyo)*	T	1643–1890	docu	Aono (2015)
Eastern Carpathians	T	1601–2004	proxy	Popa and Kern (2009)
Forfjorddalen	T	1601–2004	proxy	McCarroll et al., (2013)
Funchal (Madeira)*	R	1747–1753	inst	Alcoforado et al., (2012)
Funchal (Madeira)	T	1749–1802	inst	Alcoforado et al., (2012)
GHCN-Daily v3.27*	NR	1781–2004	inst	Menne et al., (2012)
GHCN-Monthly v2*	R	1697–2004	inst	Lawrimore et al., (2011)
GHCN-Monthly v2	P	1755–2002	inst	Peterson and Vose (1997)
Gibraltar*	R	1812–2004	inst	Wheeler (2007)
Gibraltar	T	1821–2003	inst	Wheeler (2006)
Gordon Castle	P	1782–1827	inst	Buchan (1880)
gridded MXD	T	1601–2004	proxy	Briffa et al., (2004)
HISTALP	T	1760–2004	inst	Auer et al., (2007), Böhm et al., (2010)
HISTALP	P	1763–2004	inst	Auer et al., (2007)
HISTALP*	R	1800–2004	inst	Auer et al., (2007)
Ilulissat	T	1835–2004	inst	Vinther et al., (2006)
Ireland*	R	1711–2004	inst	Murphy et al., (2018)
ISTI v1.1.0*	T	1701–2004	inst	Rennie et al., (2014)
Jämtland	T	1601–2004	proxy	Gunnarson et al., (2011)
Kawanishi-Yamagata*	T	1830–1980	docu	Ohba et al., (2013)
Kiev	T	1812–2001	inst	Brohan et al., (2006)
Kyoto*	T	1601–2004	docu	Aono and Kazui (2008)
Kyoto*	T	1601–1995	docu	Aono and Tani (2014)
Lapland	T	1601–1998	proxy	Helama et al., (2009)
London	P	1692–2004	inst	Cornes et al., (2012b)
MXD records	T	1601–2004	proxy	WSL (2019)
Nagasaki	P	1828–2000	inst	Können et al., (2003)
Nagasaki	T	1819–2000	inst	Können et al., (2003)
Naples*	T	1821–1846	inst	Dove (1847)
Nuuk	T	1784–2004	inst	Vinther et al., (2006)
Nyzhny Tagil	P	1840–1865	inst	KNMI (1871)
northern Fennoscandia*	T	1693–2004	docu	Loader et al., (2011)

(Continues)

TABLE 2 (Continued)

Data	Variable	Time	Data source	Reference
N-TREND*	T	1601–2004	proxy	Wilson et al., (2016)
PAGES2k-2017*	T, R	1601–2004	proxy	Emile-Geay et al., (2017)
Paris*	T	1658–2004	inst	Rousseau (2009, 2013)
Paris	P	1670–2004	inst	Cornes et al., (2012a)
Paris*	R	1688–2004	inst	Slonosky (2002)
Poland	T	1601–1696	docu	Przybylak et al., (2005)
Pressure collection	P	1749–2004	inst	Küttel et al., (2010)
Pyrenees	T	1601–2004	proxy	Büntgen et al., (2008)
Pyrenees	T	1601–2004	proxy	Linán et al., (2012)
Qaqortoq	T	1807–2004	inst	Vinther et al., (2006)
Rio de Janeiro*	T	1832–1843	inst	Dove (1847)
Rio de Janeiro*	R	1781–1788	inst	Farrona et al., (2012)
Rio de Janeiro	T	1781–1788	inst	Farrona et al., (2012)
Salem (Massachusetts)	P	1786–1820	inst	Van Der Schrier and Jones (2008)
San Fernando (Cadiz)*	R	1805–1944	inst	Barriendos et al., (2002)
San Fernando (Cadiz)	T	1787–1996	inst	Barriendos et al., (2002)
Savanna-la-Mar (Jamaica)*	R	1760–1786	inst	Chenoweth (2003)
Savanna-la-Mar (Jamaica)	T	1764–1776	inst	Chenoweth (2003)
Stockholm	P	1756–2004	inst	Moberg et al., (2002)
St. Lawrence Valley*	T	1742–2004	inst	Slonosky (2014)
St. Petersburg	T	1743–2004	inst	Jones and Lister (2002)
Switzerland	T	1601–1816	docu	Pfister (1999)
Switzerland	T	1601–2004	proxy	Büntgen et al., (2006)
Tatra	T	1601–2004	proxy	Büntgen et al., (2013)
Tokyo (Hachioji)*	T	1721–1941	docu	Mikami (1996)
Tokyo	T, P	1839–2000	inst	Zaiki et al., (2006)
Tomsk*	T	1830–1838	inst	Dove (1847)
Torneträsk	T	1601–2004	proxy	Grudd (2008)
Tree and coral records*	T, R	1601–2004	proxy	Neukom et al., (2014)
TRW collection*	T, R	1601–2004	proxy	Breitenmoser et al., (2014)
Tyrol	T	1601–2003	proxy	Esper et al., (2007b)
Uppsala	T	1722–2000	inst	Bergström and Moberg (2002)
Yokohama	T	1860–2000	inst	Zaiki et al., (2006)
Zafra*	T	1750–1839	docu	Fernández-Fernández et al., (2014)

Note: In the Data column, the location of single time series as well as data compilations used in the input file are presented. In the Variable column, it is indicated which variable(s) the observation represent, where NR denotes number of wet days, P is pressure, R is precipitation, and T is temperature. In the Time column, the period relevant for the assimilation is shown. The data can come from documentary (docu), instrumental (inst) and proxy sources. A reference for each time series or data collection is given in the last column.

*Denotes newly added time series /databases in EKF400v2.

additional individual records were included (Franke et al., 2017). In EKF400v2, the pressure time series were completed with one additional station from the United States (Dupigny-Giroux et al., 2007). The pressure input file contains 141 time series.

An overview of the number of records per year and the spatial distribution of observation types are shown in Figures 1 and 2. The vast majority of the data are tree-ring

width and density records throughout the reconstruction period. In Europe, a few long instrumental measurement series are available from the second half of the 17th century. In the 20th century, there are around 6,000 data entries. However, the input data file differs from the assimilated data because the data have to pass several preprocessing steps before they get assimilated (for more details, see Section 3).

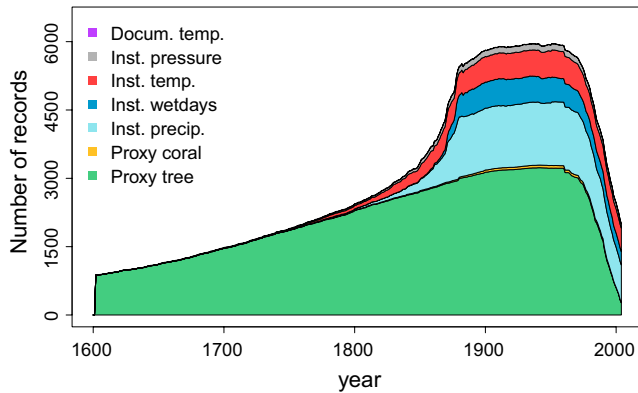


FIGURE 1 Summary of different input data types and variables per year. The documentary temperature (Docum. temp.) series (purple line) is hardly visible because only 12 series are included in the input file (see Table 2)

2.2 | Ensemble Kalman fitting

The ensemble Kalman fitting method (EKF) uses a transient offline data assimilation approach. In this offline approach, an ensemble of precomputed transient model simulations, which follow reconstructed forcings and boundary conditions, are combined with the observations. In our case, the CCC400 model simulations are blended with the above described observational data. Applying an offline scheme is a common practice for paleoclimate reconstruction because as argued in Bhend et al., (2012) predictability comes from the external forcings and boundary condition prescribed to the model simulations. Moreover, the offline approach allows simple testing of changes applied to the method since the simulations do not have to be recalculated.

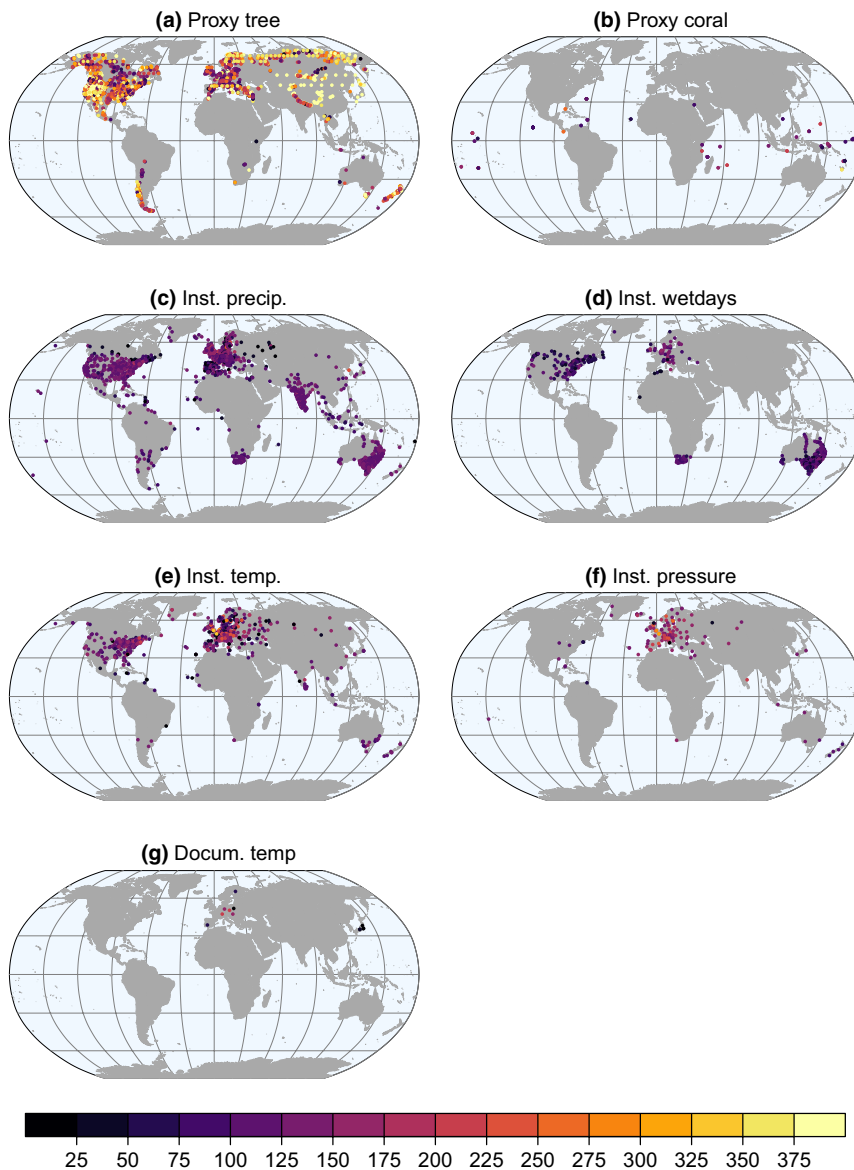


FIGURE 2 Spatial distribution of the input data types. The colours indicate the length of each series in years

The model and observational information are combined in the EKF following the procedure proposed by Whitaker and Hamill (2002); the ensemble mean (\bar{x}) and the ensemble members, transformed to deviations from the mean ($x'_i = x_i - \bar{x}$), are updated separately:

$$\bar{x}^a = \bar{x}^b + \mathbf{K}(y - H\bar{x}^b) \quad (1)$$

$$x'^a = x'^b + \tilde{\mathbf{K}}(Hx'^b) \quad (2)$$

where x^a denotes the analysis and x^b is the background state vector, obtained from the model simulation. y represents the observations. \mathbf{H} is the forward operator, that connects the model space with the observation space. In our setup, \mathbf{H} depends on the assimilated variable but it is always linear. \mathbf{K} and $\tilde{\mathbf{K}}$ are the Kalman gain matrix and the reduced Kalman gain matrix:

$$\mathbf{K} = \mathbf{P}^b \mathbf{H}^T (\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R})^{-1} \quad (3)$$

$$\tilde{\mathbf{K}} = \mathbf{P}^b \mathbf{H}^T ((\sqrt{\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R}})^{-1})^T \times (\sqrt{\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R}} + \sqrt{\mathbf{R}})^{-1} \quad (4)$$

The Kalman gain matrix based on the error estimates of the model (\mathbf{P}^b) and of the observation (\mathbf{R}) determines how to incorporate the observation into the model. \mathbf{P}^b is the background-error covariance matrix, which in the case of ensemble-based methods is calculated from the perturbations of the ensemble members. \mathbf{R} is the observation-error covariance matrix, and it is assumed that its elements are uncorrelated.

The EKF is implemented to process observations serially, which simplifies the assimilation process (Whitaker & Hamill, 2002). Due to the small ensemble size (30 members), the background-error covariance matrix is localized in order to prevent long-range spurious correlations. We apply a localization method, in which \mathbf{P}^b is multiplied element-wise with a distant-dependent correlation function:

$$G = \exp\left(-\frac{z^2}{2L^2}\right), \quad (5)$$

where z is the distance in km between two grid boxes and L is the length-scale parameter. The length-scale parameters used in the localization were calculated from the CCC400 model simulations for each variable (see Table 1 in Franke et al., 2017). The cross-covariances are localized by applying the smaller L values of the two variables in Equation (5).

In the newly generated climate reconstruction, \mathbf{P}^b is replaced with $\mathbf{P}^{\text{blend}}$ to better represent the covariances (for more details see Section 3).

3 | OVERVIEW OF THE ASSIMILATION PROCESS

In the production of the new reconstruction (EKF400v2), the basic features are consistent with the experimental design of EKF400v1. The assimilation is conducted on anomaly level. Hence, both the model simulations and the observations are transformed to anomalies calculated from a 71-year running climatology of the model and the observations. The climatology used for calculating the anomaly is shorter at the ends of the reconstructed period. We chose this procedure because the multi-decadal to centennial-scale variability may not be preserved in many proxy records in the collections used here (e.g. Franke et al., 2013; Tardif et al., 2019). This way low frequency variability is consistent with the model response to external forcings. The assimilation window is 6 months long, lasting from October to March and from April to September. The assimilation window was chosen in accordance with the growing season of trees, to assist the assimilation of time-averaged proxy records. Therefore, x^b contains 6 months data of variables of interest.

The observational network is substantially extended by including new datasets and newly digitized data, as well as adding new data types such as precipitation time series and coral proxy records to the assimilated data. The observations are quality checked before they are assimilated as in EKF400v1. Proxy records and instrumental measurements are excluded if they are more than five standard deviations away from their mean in the current time step.

Although proxy records are still assimilated by creating a forward operator using a multiple regression technique in EKF400v2, several modifications have been made. Tree-ring series are calibrated between the 1901-1971 period. In the new version, temperature-sensitive proxy series are calibrated with the Berkeley Earth surface temperature data (Rohde et al., 2013) instead of the Climatic Research Unit's gridded time-series version 3.10 dataset (CRU TS3.10, Harris et al., 2014) because of ocean coverage and higher spatial resolution. The MXD proxies are removed from the PAGES2k-2017 dataset because they are already included in the other MXD collections. The remaining tree-ring proxy records in the PAGES2k-2017 and TRW collection of Breitenmoser et al., (2014) can be both temperature- and precipitation-sensitive (Franke et al., 2020a). Therefore, to relate the proxy signal with instrumental measurements, for precipitation calibration the CRU TS4.03 (Harris & Jones, 2019) and for temperature calibration the Berkeley Earth surface temperature data are used. The regression model is built from 1 to 6 subsequent months of temperature variables in the case of temperature-sensitive proxies and from both temperature and precipitation variables in the case of temperature- and precipitation-sensitive proxies. Hence, a maximum of 12

explanatory variables are included in the multiple linear regression model, which would theoretically be the case when a tree would respond to temperature and precipitation during the entire growing season. (In our multiple linear regression models, we found no considerable multicollinearity among the explanatory variables, which could lead to less stable models.) The optimal way of assimilating various tree-ring databases was investigated in a previous study (Franke et al., 2020a). Based on the results of the experiments, the most skilful reconstruction was obtained when all databases were combined and the assimilation of proxy records was preceded with a careful quality check, discarding proxy records with insignificant climatic information. Therefore, following their findings proxy records are assimilated only if they contain a clear climatic signal. To assess whether a proxy record has a significant climatic signal, an F-Test (Neter et al., 1996) is conducted and the null-hypothesis (all of the regression coefficients are equal to zero) must be rejected with $p < 0.05$; otherwise, the proxy record is excluded. Furthermore, tree-ring proxies are screened to remove possible duplicates by checking for multiple records in a $0.1^\circ \times 0.1^\circ$ grid box, and if more than one proxy belongs to the same grid box, only the one with the smallest residual is kept.

In the PAGES2k-2017 database, coral oxygen-isotope and Sr/Ca signals are available at sub-annual resolution. Because we only have $\delta^{18}\text{O}$ information for all corals and only rarely additional Sr/Ca measurements, we use a simple approach and treat corals as only being temperature-sensitive. The regression coefficients for coral records are calculated using Berkeley Earth, which also provides information from oceanic regions by including the Met Office Hadley Centre sea surface temperature (HadSST) dataset (Kennedy et al., 2011). Since corals grow all year round, seasonal averages are created from the sub-annual coral data of the PAGES2k-2017 database (Emile-Geay et al., 2017) for the October-March and April-September time periods. The aggregated seasonal values are calibrated with the HadSST data between October 1901 and September 1971. Therefore, a maximum of 6 regression coefficients are obtained from the multiple linear regression, if corals respond to sea surface temperature in the entire season. Coral records are treated as surface temperature observations under the assumption that the surface water in which the corals grow is closely related to air temperature above.

Instrumental and documentary data are assimilated as described in Franke et al., (2017). The **H** operator extracts the closest grid cell to the observation location from the model simulations which is then used in the assimilation process. Instead of using a horizontal interpolation, we averaged the data per grid cell if several same type of observations belonged within one grid box (smoothing y). We estimated instrumental temperature and sea-level pressure observations errors based on the methods of Desroziers et al., (2005) and

Wartenburger et al., (2013). We found similar errors estimates as were used by Franke et al., (2017) in a previous study. The prescribed observational error of instrumental temperature and sea-level pressure measurements are $\sqrt{0.9}$ K and $\sqrt{10}$ hPa, respectively. The error in instrumental precipitation measurement is estimated as 30% of the measured data or at least 10 mm, and the estimated error in the number of wet days is 2 days (Valler et al., 2020). The assigned error of documentary data is 0.5 standard deviations. The error of proxy records is estimated with the residuals of the multiple regression fit. Possible duplicates in the instrumental data were not removed; however, the same type of instrumental observations are averaged if more than one are located in the same grid box. In the case if both documentary and instrumental data are available from the same location, only the instrumental measurement is assimilated. However, proxy records are used in all cases, because they get substantially less weight due to their larger errors.

The background-error covariance matrix in EKF400v1 is calculated from the 30 ensemble members and localized with the previously defined L values (Franke et al., 2017). The background-error covariance matrix plays a key role in the assimilation; therefore, it is essential to estimate it correctly. Valler et al. (2019) tested different methods for improving the estimation of \mathbf{P}^b . Based on those results, in EKF400v2 the background-error covariance matrix is not only estimated from the sample covariance but combined with a climatological covariance matrix (\mathbf{P}^{clim}). First, we create a climatological state vector (x^{clim}) by randomly selecting 100 years from the 405-year long 30-member ensemble runs; that is, the climatological state vector consists of 100 random atmospheric states. \mathbf{P}^{clim} is calculated from this climatological state vector (Valler et al., 2019). In the new reconstruction, x^{clim} is recreated every year from 100 randomly selected members. The two covariance matrices are combined as:

$$\mathbf{P}^{\text{blend}} = \beta_1 \mathbf{P}^b + \beta_2 \mathbf{P}^{\text{clim}}, \quad (6)$$

where β_1 and β_2 are the weights given to the covariance matrices and $\mathbf{P}^{\text{blend}}$ denotes the blended covariance matrix. Because \mathbf{P}^{clim} is calculated from a larger ensemble, higher L values are applied for its localization. In the reconstruction experiment, we double the L values used for localizing \mathbf{P}^{clim} and weight the two covariance matrices equally (β_1 and β_2 are 0.5). This setup has been shown to improve the skill of the reconstruction (Valler et al., 2019). Hence, EKF400v2 is generated by replacing \mathbf{P}^b with $\mathbf{P}^{\text{blend}}$ in Equations (3) and (4). Additionally, as found in previous sensitivity experiments, the reconstruction had the highest skill when both x^b and x^{clim} were updated with the observational information (Valler et al., 2019); thus, the same procedure was applied here as well. Since the order of observations may affect the reconstruction (e.g. Greybush et al., 2011), we assimilate them in the

following order: (1) instrumental measurements (temperature, pressure and finally precipitation); (2) documentary data; (3) proxy records.

In addition to the above-mentioned technical modification, the covariance matrix is also localized in time. It is a special feature of the EKF method, which is necessary due to combining several months into one assimilation window (Valler et al., 2019). Time localization is applied to monthly instrumental and documentary data by setting all cross-months covariances to zero. Therefore, these observations only influence the model fields of the current month.

4 | EVALUATING THE SKILL OF EKF400v2

4.1 | Skill analysis in the 20th century

The EKF400v2 reconstruction is compared first to the CCC400 model simulations, and later, a comparison between version 1 and version 2 of EKF400 is also shown. A leave-one-out validation was already carried out for version 1, and this is one reason why we show the improvements over version 1. We investigate the skill of the analysis ensemble mean (\bar{x}^a) by using the Berkeley Earth surface temperature data (Rohde et al., 2013), the CRU TS4.03 (Harris & Jones, 2019) and the Hadley Centre's monthly historical mean sea-level pressure (HadSLP2, Allan & Ansell, 2006) datasets as references for temperature, precipitation and sea-level pressure evaluation, respectively. The EKF400v2 is not fully independent from these data products mainly because of the assimilated instrumental measurements which start before

1880 and may have been included in the gridded instrumental datasets as well. Regression coefficients of the proxy forward models are estimated from the Berkeley Earth and the CRU TS4.03 datasets, which are used to map model data to proxy records; the errors of proxy records are estimated with the regression residuals. As a consequence, the absolute skill of the reconstruction is likely overestimated. Hence, we also compare the skill to the previous version to better capture the improvements and use newly digitized temperature series for independent evaluation.

Although our reconstruction has monthly resolution, the skill metrics shown in this Section depict seasonal averages calculated over the October–March (winter) and April–September (summer) time periods between 1902 and 2002. Figure 3 shows the correlation coefficient differences between the correlations calculated from the analysis and the reference data and the background state vector and the reference data. Correlation differences are shown because the forced model simulations already correlate with instrumental observations. Correlation differences are almost always positive for all the three variables in both half-years. The largest improvement in the temperature field is over the Northern Hemisphere, where most of the observations are located. The improvement over Europe, especially in summer (Figure 3b) and over India in winter (Figure 3a) is smaller, because the model already performs relatively well in these regions (Figure S1). Correlation increases over the central Pacific Ocean, presumably due to the assimilation of coral records. Notable increases in the correlation coefficients of precipitation are seen in Figure 3c and d. The increase is more concentrated around the instrumental precipitation measurements. The larger increase is also due to the relatively worse performance of the model

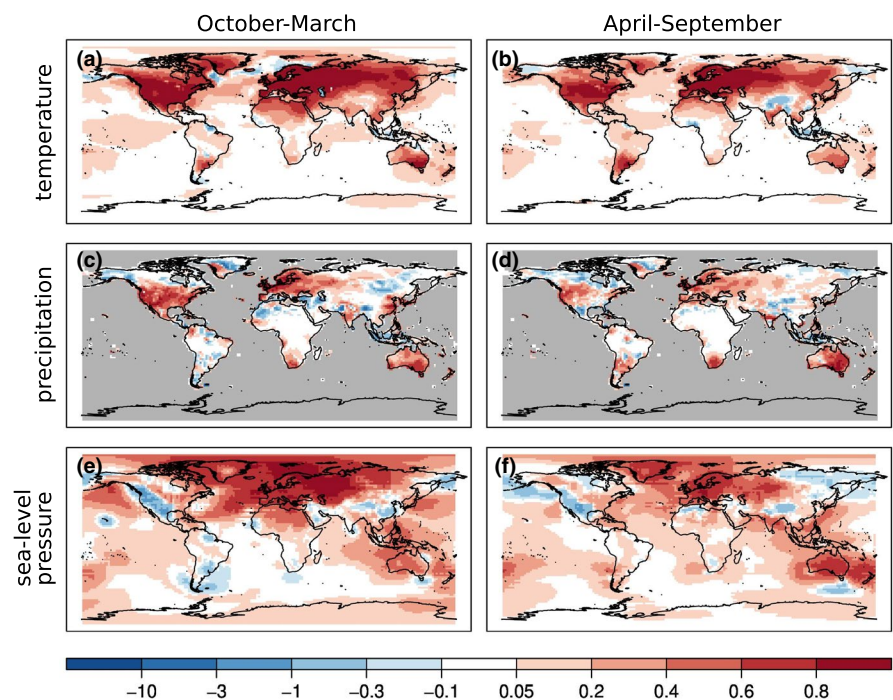


FIGURE 3 Spatial distribution of RMSESS values of temperature (a, b), precipitation (c, d) and sea-level pressure (e, f) in the two seasons. The RMSESS is calculated using the ensemble mean of the analysis and the model simulation as well as gridded instrumental data as reference. The grey shaded areas indicate the region where no reference data are available

in simulating precipitation accurately. The sea-level pressure reconstruction shows improvement mainly over the North Atlantic region and Australia, where plenty of pressure measurements are available.

Figure 4 shows the reconstruction skill in terms of the root mean square skill score (RMSESS, Wilks, 2011). While correlation expresses the covariability between two time series, RMSESS also rewards correct estimation of the amplitude. RMSESS is calculated as:

$$\text{RMSESS} = 1 - \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i^u - x_i^{\text{ref}})^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i^f - x_i^{\text{ref}})^2}}, \quad (7)$$

where in our case, x^u is the ensemble mean of the reconstruction (\bar{x}^b), x^f is the ensemble mean of the model simulation (\bar{x}^b), x^{ref} is the reference dataset, and i refers to the time step. The RMSESS is calculated from the 71-year anomalies. Positive RMSESS values indicates improvement over the CCC400 model simulation ensemble mean, while RMSESS = 1 means that the reconstruction is equal to the reference data (a perfect reconstruction). In previous publications, the RMSESS was termed reduction of error (RE; e.g. Franke et al., 2017).

The RMSESS values of temperature are mainly positive (Figure 4a, b). As expected, the largest improvement is over the land areas in the Northern Hemisphere, but the reconstruction also shows some skill over oceanic areas. The skill of the reconstruction is very similar in both seasons in the 20th century when instrumental data are available.

The RMSESS values of precipitation depict more spatial variability (Figure 4c, d). Over Europe, India, Australia and South Africa, the skill is positive both in winter and in summer. The precipitation reconstruction performs better over North and Central America in the October-March half-year. Similarly, the reconstruction is more skilful in the April-September period over the southern part of South America. In the sea-level pressure reconstruction, the largest improvement occurred over the North Atlantic sector and Eurasia, with more notable improvement in the October-March period. In the Southern Hemisphere, the improvement is also more pronounced during winter time from April to September, but the degree of improvement is smaller. However, negative RMSESS values characterize, for example, the North American Cordillera area in both half-years.

To investigate the improvement of version 2 in comparison to version 1, the RMSESS of both versions are calculated and their differences (v2 minus v1) are shown in Figure 5. Since EKF400v1 was generated by using every second grid box in both the latitude and longitude, the comparison was made on the resolution of EKF400v1.

The reconstructed temperature field already had very high RMSESS values over the Northern Hemisphere in EKF400v1 (see Figure 8 in Franke et al., 2017), but even in this region further improvements can be seen (Figure 5a, b). Large differences in RMSESS values of the precipitation occur between the two versions, especially over the extratropical Northern Hemisphere in the October-March period (Figure 5c). However, the skill is not always positive, for example over North Africa in winter or over North India and the Tibetan plateau in summer. The reconstructed sea-level pressure field

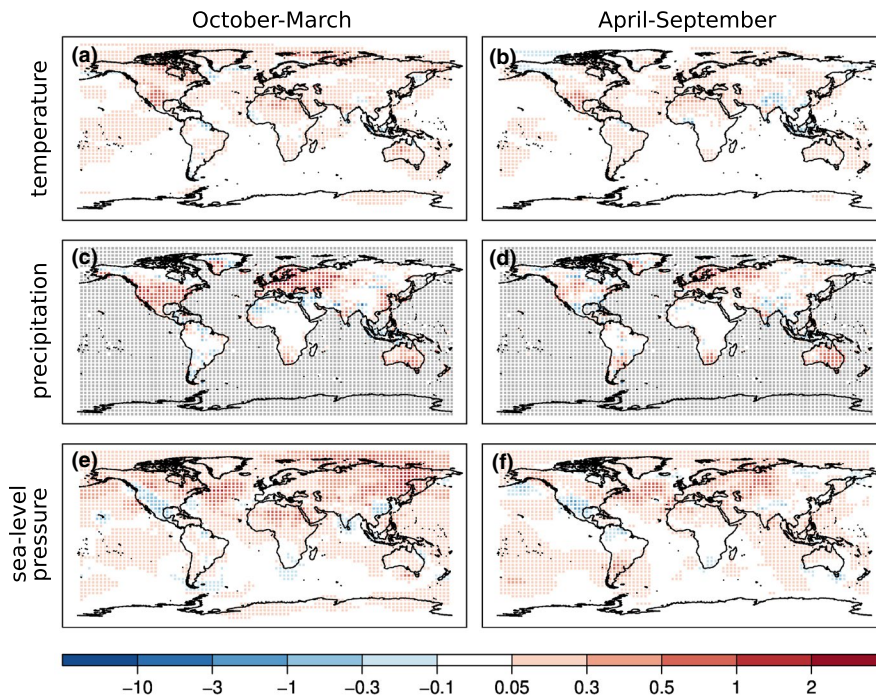
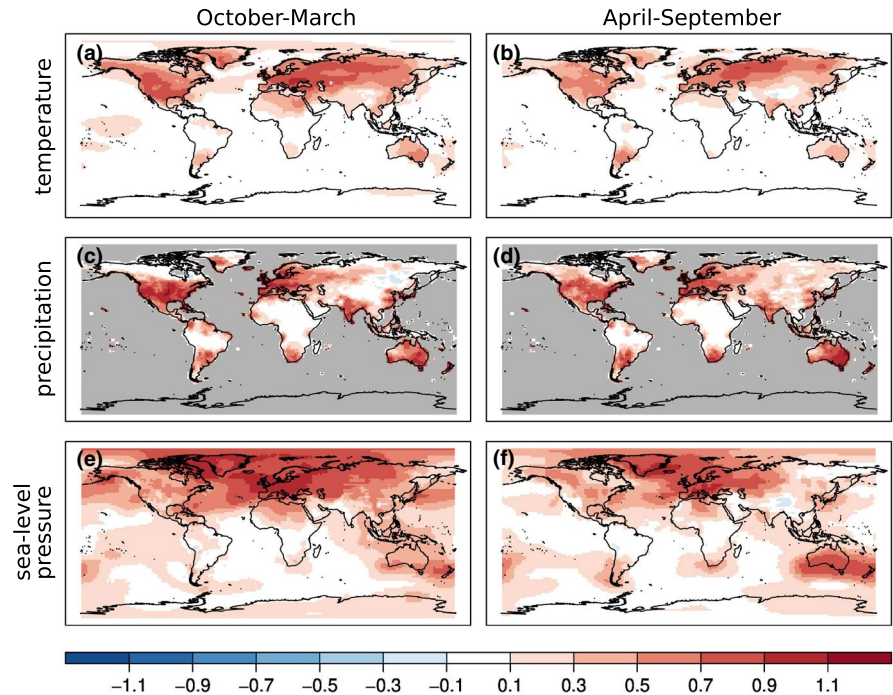


FIGURE 4 Spatial distribution of RMSESS values differences between EKF400v2 and EKF400v1 (EKF400v2 minus EKF400v1). The differences are calculated on the resolution of EKF400v1 for temperature (a, b), precipitation (c, d) and sea-level pressure (e, f) in the two seasons. Positive values imply that v2 has higher skill than v1. The grey shaded areas indicate the region where no reference data are available. Note that the colour scale is nonlinear and asymmetric

FIGURE 5 Spatial distribution of correlation coefficients differences between the analysis mean of EKF400v2 and the ensemble mean of the forced simulations of temperature (a, b), precipitation (c, d) and sea-level pressure (e, f) in the two seasons. The grey shaded areas indicate the region where no reference data are available



also mainly exhibits improvements over EKF400v1. Large areas both over land and ocean are characterized by positive RMSESS values.

Prior to the 20th century, the skill of reconstructed temperature field was evaluated with newly digitized data mainly from the 19th century (Dove, 1847; Weselowskij, 1857), which will be assimilated in the next version. Given that most of the new data come from coastal stations, we compared the independent time series with the analysis ensemble mean of the closest grid cell that we considered sufficiently representative of land surface temperature and that is present both in v1 and v2; for this reason, the selected grid cell can be up to 200 kilometres away from the location of the observations. The data at high northern latitudes show strong correlation with EKF400v1 but they further increased with EKF400v2 (Figure 6c,d,g,i). In the low latitudes, some improvement can still be seen in EKF400v2 compared to EKF400v1, although the uncertainty of the reconstruction there remains rather large (Figure 6e,f,h). In the cases of the two examples from the Southern Hemisphere, no nearby observations were assimilated in EKF400v1 over the analysed periods and no correlation is found with the independent temperature measurements (Figure 6a,b). In EKF400v2, tree-ring proxies from the Southern Hemisphere were added to the assimilated observations which can affect the growing season climate fields; indeed, we can see some increase in the correlation between the Auckland temperature series and EKF400v2 (Figure 6a). However, no improvements appear for Cape Town (Figure 6b). Tree growth in South Africa is mostly limited by moisture and not temperature which could explain why the reconstructed temperature of EKF400v2 has not been improved by the assimilation.

4.2 | Case studies

In order to further evaluate the skill of EKF400v2 in the pre-20th century time, two case studies are investigated. We chose one summer- and one wintertime case study. In the summer-time case study, we analyse a drought period over central Europe in the early 18th century and evaluate the reconstruction skill with independent early instrumental measurements that were not assimilated. In the wintertime case study, we look at known large-scale teleconnection patterns by comparing the 19th-century patterns with their known 20th-century counterpart under the assumption of stationarity.

4.2.1 | Summertime example: Central European drought of 1726-1728

According to documentary evidence, the late 1720s coincide with one of the driest periods in recorded history in Central Europe. Brázdil et al., (2013), for instance, reconstructed drought periods in the Czech Lands between 1501-1804 and found the 1720s to be the decade with the most intense drought over the whole period. In 1726, in particular, six consecutive 'dry' months (April to September) were described in historical documents, although the impact on the population was more evident in 1727 (only June not classified as dry) and 1728 (whole summer 'very dry') with famine and even a rare locust plague (Brázdil et al., 2019).

This is also a decade in which early instrumental observations were carried out in several European cities, although only few of them have been recovered (Brönnimann et al., 2019a). Here we use the pressure data of Uppsala (Bergstrom

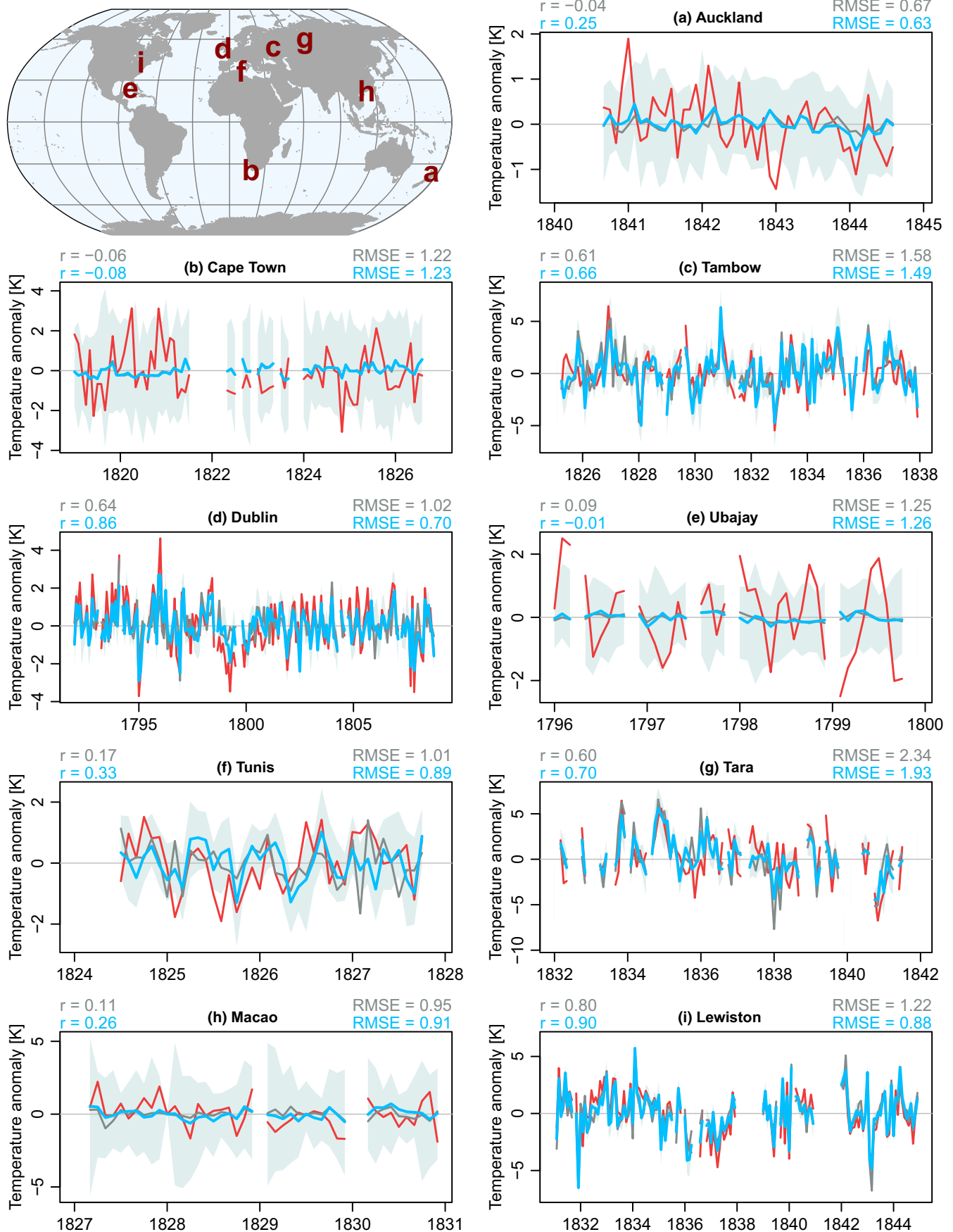


FIGURE 6 Monthly temperature anomalies of the independent observations (red), the analysis ensemble mean of EKF400v1 (grey) and EKF400v2 (blue). With light blue, the ensemble range of EKF400v2 is shown. The correlations (r) and the root mean square error (RMSE) between the reconstructions and the independent measurements are given on the figures. On the map, the location of observations is presented

& Moberg, 2002) and Padua (Camuffo et al., 2006), which have not been assimilated in either version of EKF400, to validate our reconstruction of the May to September periods in 1726, 1727 and 1728. The locations of the two stations are shown in Figure 7.

Figures 7 and 8 show the reconstructed 500 hPa geopotential height and relative anomalies of precipitation fields over Europe in EKF400v2 for each month (equivalent figures for EKF400v1 are provided in the Supplement (Figures S2, S3)). As expected, most panels show negative precipitation anomalies in Central Europe, often reaching -60% . The anomalies of the geopotential height indicate a similar pattern in May in all three years, with an anomalous ridge over the continent. Weaker positive anomalies are reconstructed for most of Europe in nearly all months with the exception of September 1726 and 1728. These results clearly suggest drought conditions.

The instrumental pressure records are compared with EKF400v1 and EKF400v2 in Figure 9. The correlation between observations and reconstruction is markedly higher for both stations in the new version. In both versions, only one pressure record (London, UK) is assimilated in that period. The skill increase can be due to the assimilation of precipitation as well as the additional tree-ring width records and methodological upgrades.

4.2.2 | Wintertime example: La Niña impacts in the 19th century

The El-Niño Southern Oscillation (ENSO) is known to have significant impacts on extratropical circulation (e.g. Livezey et al., 1997). In the mid-latitudes of the Northern Hemisphere, many of the impacts are largest in late winter (January to March, or JFM) (e.g. Brönnimann et al., 2007; Schmidt et al., 2001). In order to better capture the ENSO signal (in this case, La Niña), a composite of all La Niña years in the 19th century (1801, 1820, 1822, 1835, 1842, 1847, 1863, 1872, 1887, 1890, 1893 (based on Brönnimann et al., 2007)) was calculated. In EKF400v1, the reconstruction skill for winter months in the 19th century was greatly limited by the number of observational and documentary records, particularly outside of Europe. As shown in Figure 10, the ability of the reconstruction to reproduce the La Niña signal in North America was practically equivalent to the ability of the model simulation. In version 2, numerous precipitation and temperature series are added in North America thanks to the assimilation of the GHCN-Monthly v2 precipitation and ISTI datasets (see Table 2). This brings a clear improvement in the reconstruction of the ENSO signal in the south-eastern United States, where La Niña conditions are related to strong negative

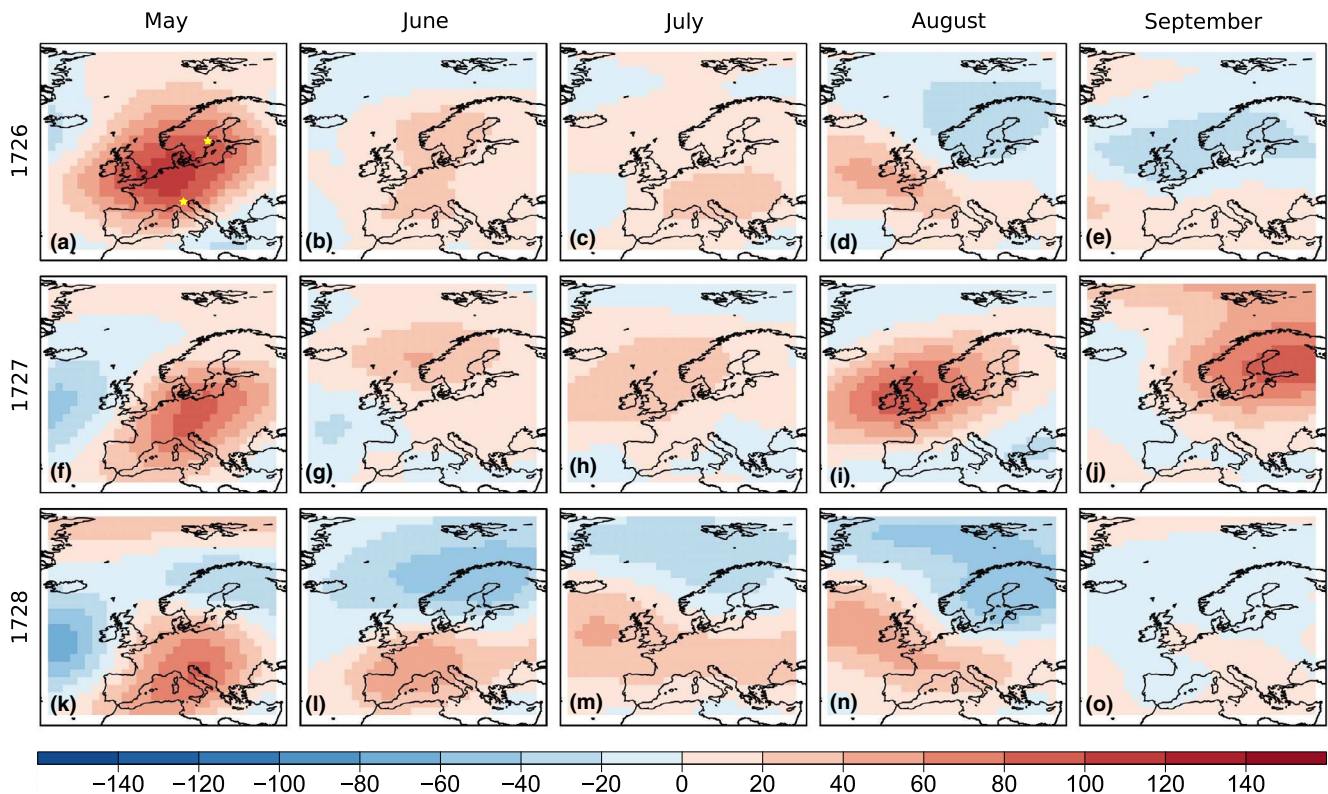


FIGURE 7 Monthly anomalies of 500 hPa geopotential height from May to September over Europe for the dry summer years between 1726 and 1728 in EKF400v2. The yellow stars indicate the location of Padua and Uppsala

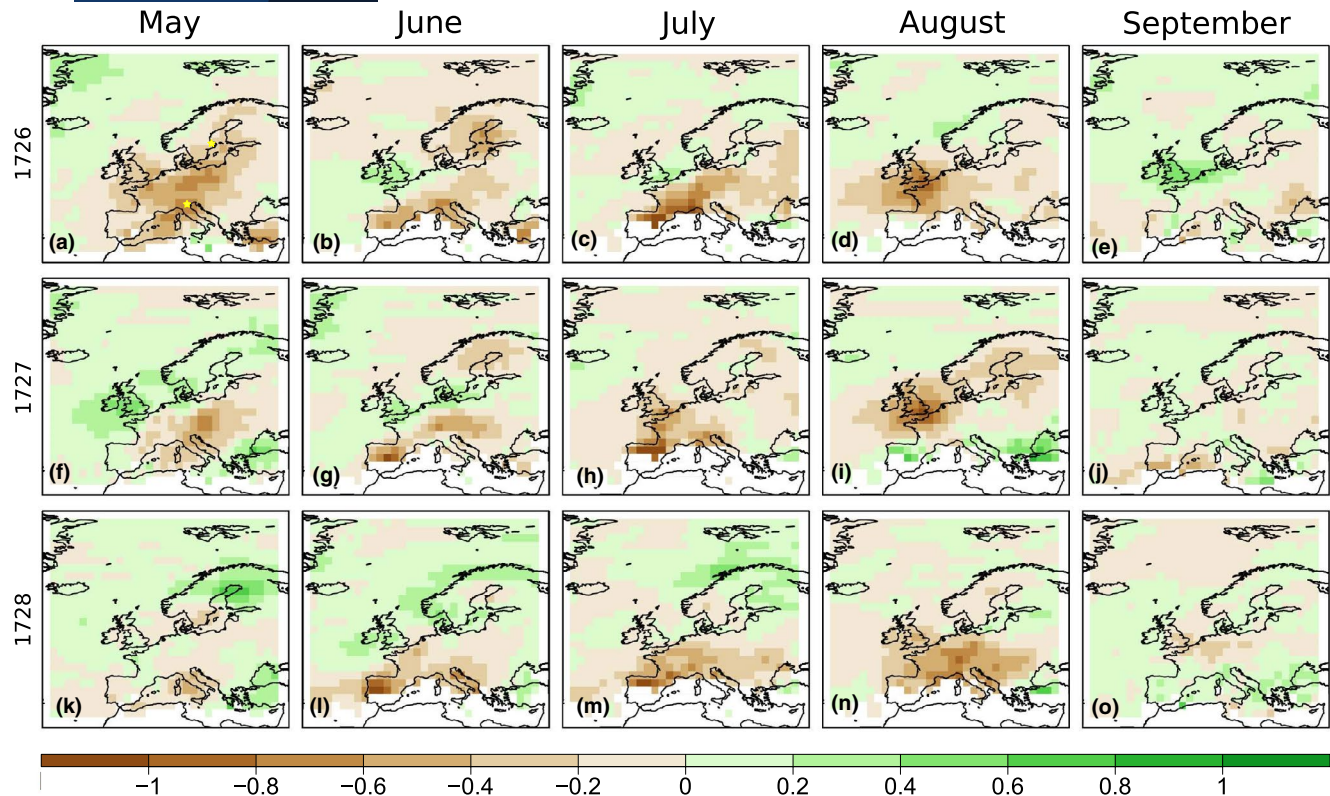


FIGURE 8 Monthly relative anomalies of precipitation from May to September over Europe for the dry summer years between 1726 and 1728 in EKF400v2. The areas left blank in the Mediterranean are regions where monthly precipitation amount is less than 10 mm in the climatology

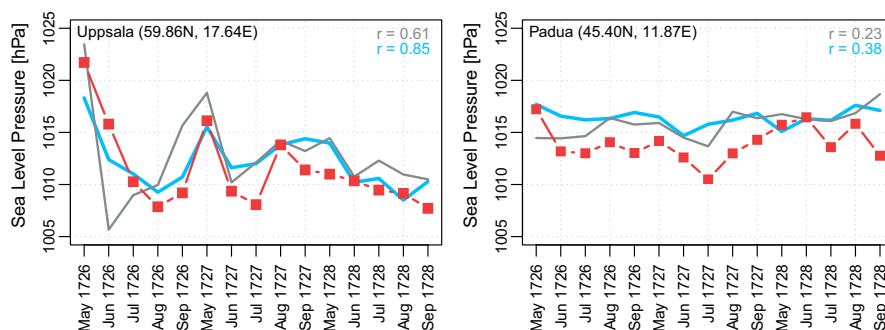


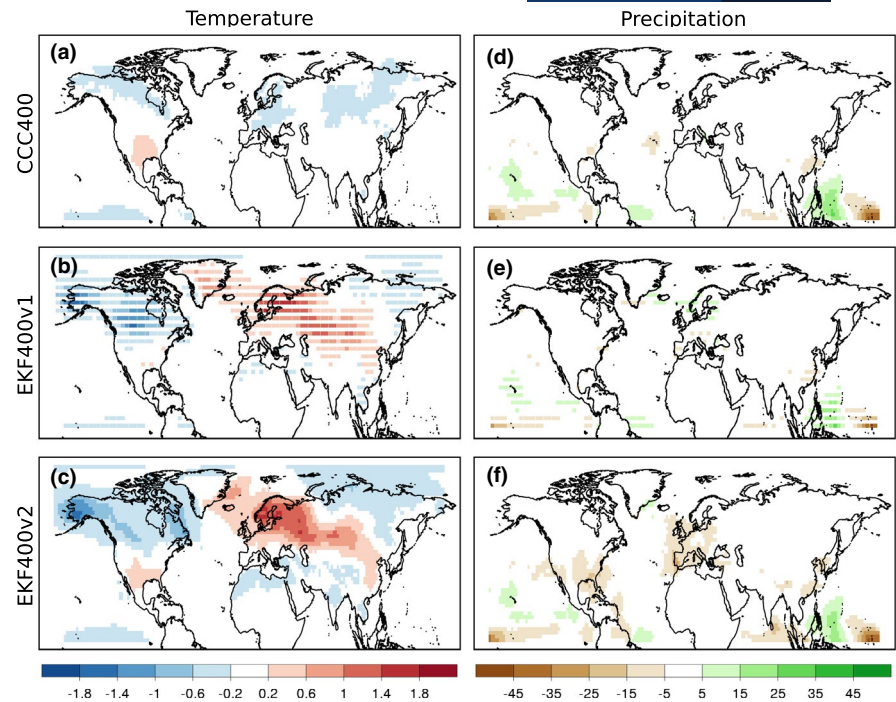
FIGURE 9 Monthly time series comparison among EKF400v1 (grey), EKF400v2 (blue) and instrumental pressure measurements (red) between 1726 and 1728. The correlations between the reconstructions and the independent measurements are given on the figures

precipitation anomalies (Schmidt et al., 2001). Assuming the stationarity of ENSO impacts on centennial scale (the La Niña composites of temperature and precipitation in the 20th century calculated from EKF400v2 are shown in Figure S4), positive temperature anomalies in the northern Gulf of Mexico (Livezey et al., 1997) are also more realistically reconstructed. ENSO also has an impact on the climate of Europe, which in the case of a La Niña event resembles to the positive North Atlantic Oscillation signal in Europe (Brönnimann et al., 2007). This signal was already well captured in version 1, although negative precipitation anomalies over Southern and Western Europe are more prominent in version 2.

5 | SUMMARY

EKF400v2 was generated on the full spatial resolution of the model simulations, therefore it provides higher spatial resolution than EKF400v1. Over the 1902-2002 period, generally high skill characterizes the EKF400v2 reconstruction but skill varies with the reconstructed variables, seasons and regions. An overall improvement in the skill of the three key variables (temperature, precipitation, sea-level pressure) between the two versions has been found. In the early period, when primarily proxy data are available, most information is added to the summer growing season. To assess the skill prior to the twentieth century, we have looked at two case studies using

FIGURE 10 Composite figure of temperature anomaly [K] and precipitation anomaly [mm], calculated from the La Niña years in the 19th century (1801, 1820, 1822, 1835, 1842, 1847, 1863, 1872, 1887, 1890, 1893; based on Brönnimann et al., (2007)). The composite of ensemble mean of the model simulation (a, d), the EKF400v1 (b, e) and the EKF400v2 (c, f) are shown



independent data for evaluation. In the case of the summer-time drought analysis, by investigating other unobserved variable fields of the state vector such as the 500 hPa geopotential height we can gain understanding of the climate dynamics leading to the extreme dry period. The skill of the reconstruction highly depends on the number of available observations. Further back in time, especially in the winter season, when only a few data series from Europe and Japan are assimilated, EKF400v2 is mainly driven by the model simulations.

6 | DATASET LOCATION AND USAGE

The EKF400v2 monthly climate paleo-reanalysis is archived at the World Data Center for Climate (WDCC) at DKRZ, and each ensemble member and the ensemble mean are stored in separate NetCDF files (Franke et al., 2020b). The files contain the absolute values of the following variables: air temperature at 2 m, total precipitation, sea-level pressure, geopotential height at 500 and 100 hPa, eastward and northward wind components at 850 and 200 hPa and vertical velocity at 500 hPa. Data are accessible to everybody after registration.

The preliminary results with EKF400v2 highlight areas of improvement over EKF400v1 and can be used to investigate the dynamics behind past climate changes over the last 400 years. However, users must take into account the limitations of the dataset, the skill of the reconstruction changes with the regions and with the course of time. In places and times when assimilated data are abundant, the paleo-reanalysis can

be used with more confidence than, for instance, in the beginning of the reconstruction period when only proxy records are available. Since the assimilation was conducted assimilating anomalies (from each observation and model grid cell the corresponding 71-year running monthly climatologies were subtracted), long-term variability (>70 years) rise from the forcings and boundary conditions of the models. In addition, carrying out the assimilation on the anomaly level may result in negative absolute values of total precipitation. Franke et al. (2017) point out that EKF400v1 south from 60°S should not be used in any analysis due to errors in the model simulations. This feature is inherited in v2 because the same model simulations were involved in the assimilation. Conducting the assimilation on the full resolution of the model plus applying an improved assimilation scheme (blended covariance technique) implied larger computational burden. Hence, only a group of variables could be loaded at the same time and the climatological state vector, using randomly selected members, was generated separately for each group of variables.

ACKNOWLEDGEMENTS

This research has received funding from the Swiss National Science Foundation (grant no. 162668) and the European Commission - Horizon 2020 (grant no. 787574). The CCC400 simulation was performed at the Swiss National Supercomputing Centre CSCS. We thank Angela-Maria Burgdorf for collecting additional documentary records used in the assimilation.

CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest.

OPEN PRACTICES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at 10.26050/WDCC/EKF400_v2.0 Learn more about the Open Practices badges from the Center for OpenScience: <https://osf.io/tvyxz/wiki>.

ORCID

Veronika Valler  <https://orcid.org/0000-0002-2515-1433>
Yuri Brugnara  <https://orcid.org/0000-0001-8427-0064>
Stefan Brönnimann  <https://orcid.org/0000-0001-9502-7991>

REFERENCES

- Alcoforado, M.J., Vaquero, J., Trigo, R. & Taborda, J. (2012) Early Portuguese meteorological measurements (18th century). *Climate of the Past*, 8, 353–371. <https://doi.org/10.5194/cp-8-353-2012>
- Allan, R. & Ansell, T. (2006) A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850–2004. *Journal of Climate*, 19, 5816–5842. <https://doi.org/10.1175/JCLI3937.1>
- Aono, Y. (2015) Cherry blossom phenological data since the seventeenth century for Edo (Tokyo), Japan, and their application to estimation of March temperatures. *International Journal of Biometeorology*, 59, 427–434. <https://doi.org/10.1007/s00484-014-0854-0>
- Aono, Y. & Kazui, K. (2008) Phenological data series of cherry tree flowering in Kyoto, Japan, and its application to reconstruction of springtime temperatures since the 9th century. *International Journal of Climatology*, 28, 905–914. Wiley Online Library. <https://doi.org/10.1002/joc.1594>
- Aono, Y. & Tani, A. (2014) Autumn temperature deduced from historical records of autumn tints phenology of maple tree in Kyoto, Japan. *Climate in Biosphere* (in Japanese with English abstract), 14, 18–28. <https://doi.org/10.2480/cib.J-14-023>
- Auer, I., Böhm, R., Jurkovic, A., Lipa, W., Orlik, A., Potzmann, R. et al. (2007) HISTALP historical instrumental climatological surface time series of the Greater Alpine Region. *International Journal of Climatology*, 27, 17–46. <https://doi.org/10.1002/joc.1377>
- Barriendos, M., Gámez, B. & Peña, J. (1997) *Old series of meteorological readings for Madrid and Barcelona (1780–1860)*. Documentary and observed characteristics, *Advances in Historical Climatology in Spain*, pp. 157–172. Oikos-Tau: Barcelona.
- Barriendos, M., Martín-Vide, J., Peña, J.C. & Rodríguez, R. (2002) Daily meteorological observations in Cadiz-San Fernando. Analysis of the documentary sources and the instrumental data content (1786–1996). *Climatic Change*, 53, 151–170. <https://doi.org/10.1023/A:1014991430122>
- Bartholy, J., Pongrácz, R. & Molnár, Z. (2004) Classification and analysis of past climate information based on historical documentary sources for the Carpathian basin. *International Journal of Climatology*, 24, 1759–1776. <https://doi.org/10.1002/joc.1106>
- Bergström, H. & Moberg, A. (2002) Daily air temperature and pressure series for Uppsala 1722–1998. *Climatic Change*, 231–252. <https://doi.org/10.1023/A:1014983229213>
- Bhend, J., Franke, J., Folini, D., Wild, M. & Brönnimann, S. (2012) An ensemble-based approach to climate reconstructions. *Climate of the Past*, 8, 963–976. <https://doi.org/10.5194/cp-8-963-2012>
- Blodgett, L. (1857) *Climatology of the United States and of the temperature latitudes of the North American continent*. Philadelphia, J. B. Lippincott and Co.
- Böhm, R., Jones, P.D., Hiebl, J., Frank, D., Brunetti, M. & Maugeri, M. (2010) The early instrumental warm-bias: a solution for long central European temperature series 1760–2007. *Climatic Change*, 101, 41–67. <https://doi.org/10.1007/s10584-009-9649-4>
- Brázdil, R., Dobrovolný, P., Trnka, M., Kotyza, O., Řezníčková, L., Valášek, H. et al. (2013) Droughts in the Czech lands, 1090–2012 AD. *Climate of the Past*, 9, 1985–2002. <https://doi.org/10.5194/cp-9-1985-2013>
- Brázdil, R., Dobrovolný, P., Trnka, M., Řezníčková, L., Dolák, L. & Kotyza, O. (2019) Extreme droughts and human responses to them: the Czech Lands in the pre-instrumental period. *Climate of the Past*, 15, 1–24. <https://doi.org/10.5194/cp-15-1-2019>
- Breitenmoser, P.D., Brönnimann, S. & Frank, D. (2014) Forward modelling 679 of tree-ring width and comparison with a global network of tree-ring chronologies. *Climate of the Past*, 10, 437–449. <https://doi.org/10.5194/cp-10-437-2014>
- Briffa, K.R., Osborn, T.J. & Schweingruber, F. (2004) Large-scale temperature inferences from tree rings: a review. *Global and Planetary Change*, 40, 11–26. [https://doi.org/10.1016/S0921-8181\(03\)00095-X](https://doi.org/10.1016/S0921-8181(03)00095-X)
- Brohan, P., Kennedy, J.J., Harris, I., Tett, S.F. & Jones, P.D. (2006) Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *Journal of Geophysical Research: Atmospheres*, 111, D12106. <https://doi.org/10.1029/2005JD006548>
- Brönnimann, S., Allan, R., Ashcroft, L., Baer, S., Barriendos, M., Brázdil, R. et al. (2019a) Unlocking pre-1850 instrumental meteorological records: A global inventory. *Bulletin of the American Meteorological Society*, 100(12), ES389–ES413. <https://doi.org/10.1175/BAMS-D-19-0040.1>
- Brönnimann, S., Franke, J., Nussbaumer, S.U., Zumbühl, H.J., Steiner, D., Trachsel, M. et al. (2019b) Last phase of the Little Ice Age forced by volcanic eruptions. *Nature Geoscience*, 12, 650–656. <https://doi.org/10.1038/s41561-019-0402-y>
- Brönnimann, S., Xoplaki, E., Casty, C., Pauling, A. & Luterbacher, J. (2007) ENSO influence on Europe during the last centuries. *Climate Dynamics*, 28, 181–197. <https://doi.org/10.1007/s00382-006-0175-z>
- Buchan, A. (1880) The meteorology of Gordon Castle. *Journal of the Scottish Meteorological Society*, 5, 59–63.
- Büntgen, U., Frank, D., Grubb, H. & Esper, J. (2008) Long-term summer temperature variations in the Pyrenees. *Climate Dynamics*, 31, 615–631. <https://doi.org/10.1007/s00382-008-0390-x>
- Büntgen, U., Frank, D.C., Nievergelt, D. & Esper, J. (2006) Summer temperature variations in the European Alps, AD 755–2004. *Journal of Climate*, 19, 5606–5623. <https://doi.org/10.1175/JCLI3917.1>
- Büntgen, U., Kyncl, T., Ginzler, C., Jacks, D.S., Esper, J., Tegel, W. et al. (2013) Filling the Eastern European gap in millennium-long temperature reconstructions. *Proceedings of the National Academy of Sciences*, 110, 1773–1778. <https://doi.org/10.1073/pnas.1211485110>
- Burgdorf, A.-M., Brönnimann, S. & Franke, J. (2019) Two types of North American droughts related to different atmospheric circulation patterns. *Climate of the Past*, 15, 2053–2065. <https://doi.org/10.5194/cp-15-2053-2019>

- Camuffo, D., Cocheo, C. & Sturaro, G. (2006) Corrections of systematic errors, data homogenisation and climatic analysis of the Padova pressure series (1725–1999). *Climatic Change*, 78, 493–514. <https://doi.org/10.1007/s10584-006-9052-3>
- Chenoweth, M. (2003) The 18th Century Climate of Jamaica: Derived from the Journals of Thomas Thistlewood, 1750–1786. *Transactions of the American Philosophical Society*, 93(2), 1–153. American Philosophical Society, Philadelphia.
- Cook, E.R., D'Arrigo, R. & Anchukaitis, K.J. *ENSO reconstructions from long tree-ring chronologies: Unifying the differences?*, talk presented at a special workshop on "Reconciling ENSO Chronologies for the Past 500 Years" held in Moorea, French Polynesia on 2–3 April, 2008.
- Cornes, R.C., Jones, P.D., Briffa, K.R. & Osborn, T.J. (2012a) A daily series of mean sea-level pressure for Paris, 1670–2007. *International Journal of Climatology*, 32, 1135–1150. <https://doi.org/10.1002/joc.2349>
- Cornes, R.C., Jones, P.D., Briffa, K.R. & Osborn, T.J. (2012b) A daily series of mean sea-level pressure for London, 1692–2007. *International Journal of Climatology*, 32, 641–656. <https://doi.org/10.1002/joc.2301>
- Crowley, T.J., Zielinski, G., Vinther, B., Udisti, R., Kreutz, K., Cole-Dai, J. & et al. (2008) Volcanism and the little ice age. *PAGES Newsletter*, 16, 22–23.
- Demarée, G., Lachaert, P.-J., Verhoeve, T. & Thoen, E. (2002) The long-term daily central Belgium temperature (CBT) series (1767–1998) and early instrumental meteorological observations in Belgium. *Climatic Change*, 53, 269–293. <https://doi.org/10.1023/A:1014931211466>
- Derham, W. (1705) A prospect of the weather, winds, and height of the mercury in the barometer, on the first day of the month; and of the whole rain of every month in the year 1703, and the beginning of 1704. Observ'd at Towneley in Lancashire, by Ri. Towneley, Esq; and at Upminster in Essex. *Philosophical Transactions of the Royal Society of London*, 24, 1878–1881. <https://doi.org/10.1098/rstl.1704.0063>
- Desroziers, G., Berre, L., Chapnik, B. & Poli, P. (2005) Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, 131, 3385–3396. <https://doi.org/10.1256/qj.05.108>
- Dobrovolný, P., Moberg, A., Brázdil, R., Pfister, C., Glaser, R., Wilson, R. et al. (2010) Monthly, seasonal and annual temperature reconstructions for Central Europe derived from documentary evidence and instrumental records since AD 1500. *Climatic Change*, 101, 69–107. <https://doi.org/10.1007/s10584-009-9724-x>
- Dove, H.W. (1847) Über die nicht periodischen Änderungen der Temperaturvertheilung auf der Oberfläche der Erde, Vierte Abhandlung, in: *Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin*, Berlin: Königliche Akademie der Wissenschaften.
- Dupigny-Giroux, L.-A., Ross, T.F., Elms, J.D., Truesdell, R. & Doty, S.R. (2007) NOAA's climate database modernization program: rescuing, archiving, and digitizing history. *Bulletin of the American Meteorological Society*, 88, 1015–1017. <https://doi.org/10.1175/BAMS-88-7-1015>
- Emile-Geay, J., McKay, N.P., Kaufman, D.S., Von Gunten, L., Wang, J., Anchukaitis, K.J. et al. (2017) A global multiproxy database for temperature reconstructions of the Common Era. *Scientific Data*, 4, 170088. <https://doi.org/10.1038/sdata.2017.88>
- Esper, J., Büntgen, U., Frank, D., Pichler, T. & Nicolussi, K. (2007b) Updating the Tyrol tree-ring dataset, TRACE-Tree Rings in Archaeology. *K. Haneca Climatology and Ecology*, 5, 80–85.
- Farrona, A., Trigo, R., Gallego, M. & Vaquero, J. (2012) The meteorological observations of Bento Sanches Dorta, Rio de Janeiro, Brazil: 1781–1788. *Climatic Change*, 115, 579–595. <https://doi.org/10.1007/s10584-012-0467-8>
- Fernández-Fernández, M., Gallego, M., Domínguez-Castro, F., Trigo, R., García, J., Vaquero, J. et al. (2014) The climate in Zafra from 1750 to 1840: history and description of weather observations. *Climatic Change*, 126, 107–118. <https://doi.org/10.1007/s10584-014-1201-5>
- Franke, J., Brönnimann, S., Bhend, J. & Brugnara, Y. (2017) A monthly global paleo-reanalysis of the atmosphere from 1600 to 2005 for studying past 785 climatic variations. *Scientific Data*, 4, 170076. <https://doi.org/10.1038/sdata.2017.76>
- Franke, J., Frank, D., Raible, C.C., Esper, J. & Brönnimann, S. (2013) Spectral biases in tree-ring climate proxies. *Nature Climate Change*, 3, 360–364. <https://doi.org/10.1038/nclimate1816>
- Franke, J., Valler, V., Brönnimann, S., Neukom, R. & Jaume-Santero, F. (2020a) The importance of input data quality and quantity in climate field reconstructions—results from the assimilation of various tree-ring collections, *Climate of the Past*, 16, 1061–1074. <https://doi.org/10.5194/cp-16-1061-2020>
- Franke, J., Valler, V., Brugnara, Y. & Brönnimann, S. (2020b) *Ensemble Kalman Fitting Paleo-Reanalysis Version 2 (EKF400 v2)*, World Data Center for Climate (WDCC) at DKRZ, https://doi.org/10.26050/WDCC/EKF400_v2.0
- Glaser, R. & Riemann, D. (2009) A thousand-year record of temperature variations for Germany and Central Europe based on documentary data. *Journal of Quaternary Science: Published for the Quaternary Research Association*, 24, 437–449. <https://doi.org/10.1002/jqs.1302>
- Greybush, S.J., Kalnay, E., Miyoshi, T., Ide, K. & Hunt, B.R. (2011) Balance and ensemble Kalman filter localization techniques. *Monthly Weather Review*, 139, 511–522. <https://doi.org/10.1175/2010MWR3328.1>
- Grudd, H. (2008) Torneträsk tree-ring width and density AD 500–2004: a test of climatic sensitivity and a new 1500-year reconstruction of north Fennoscandian summers. *Climate Dynamics*, 31, 843–857. <https://doi.org/10.1007/s00382-007-0358-2>
- Gunnarson, B.E., Linderholm, H.W. & Moberg, A. (2011) Improving a treering reconstruction from west-central Scandinavia: 900 years of warmseason temperatures. *Climate Dynamics*, 36, 97–108. <https://doi.org/10.1007/s00382-010-0783-5>
- Hakim, G., Annan, J., Brönnimann, S., Crucifix, M., Edwards, T., Goosse, H. et al. (2013) Overview of data assimilation methods. *PAGES Newsletter*, 21, 72–73. <https://doi.org/10.7892/boris.39169>
- Harris, I. & Jones, P. (2019) *CRU TS4.03: Climatic Research Unit (CRU) Time-Series (TS) version 4.03 of high-resolution gridded data of month-by-month variation in climate (Jan. 1901–Dec. 2018)*. Centre for Environmental Data Analysis, <https://catalogue.ceda.ac.uk/uuid/10d3e3640f004c578403419aac167d82>, accessed 8 Dec 2019
- Harris, I., Jones, P.D., Osborn, T.J. & Lister, D.H. (2014) Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *International Journal of Climatology*, 34(3), 623–642. <https://doi.org/10.1002/joc.3711>
- Helama, S., Timonen, M., Holopainen, J., Ogurtsov, M.G., Mielikäinen, K., Eronen, M. et al. (2009) Summer temperature variations in Lapland during the Medieval Warm Period and the Little Ice Age relative to natural instability of thermohaline circulation on multi-decadal and multi-centennial scales. *Journal of Quaternary Science: Published for the Quaternary Research Association*, 24, 450–456. <https://doi.org/10.1002/jqs.1291>

- Jones, P.D. & Lister, D.H. (2002) The Daily Temperature Record for St. Petersburg (1743–1996). *Climatic Change*, 53, 253–267. <https://doi.org/10.1023/A:1014918808741>
- Kennedy, J., Rayner, N., Smith, R., Parker, D. & Saunby, M. (2011) Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 1. Measurement and sampling uncertainties. *Journal of Geophysical Research: Atmospheres*, 116, D14104. <https://doi.org/10.1029/2010JD015220>
- KNMI (1871) *Nederlandsch Meteorologisch Jaarboek voor 1870*, Kemink en Zoon.
- Koch, D., Jacob, D., Tegen, I., Rind, D. & Chin, M. (1999) Tropospheric sulfur simulation and sulfate direct radiative forcing in the Goddard Institute for Space Studies general circulation model. *Journal of Geophysical Research: Atmospheres*, 104, 23799–23822. <https://doi.org/10.1029/1999JD900248>
- Können, G.P., Zaiki, M., Baede, A.P.M., Mikami, T., Jones, P.D. & Tsukahara, T. (2003) Pre-1872 extension of the Japanese instrumental meteorological observation series back to 1819. *Journal of Climate*, 16, 118–131. [https://doi.org/10.1175/1520-0442\(2003\)016<0118:PEOTJI>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<0118:PEOTJI>2.0.CO;2)
- Küttel, M., Xoplaki, E., Gallego, D., Luterbacher, J., Garcia-Herrera, R., Allan, R. et al. (2010) The importance of ship log data: reconstructing North Atlantic, European and Mediterranean sea level pressure fields back to 1750. *Climate Dynamics*, 34, 1115–1128. <https://doi.org/10.1007/s00382-009-0577-9>
- Lawrimore, J.H., Menne, M.J., Gleason, B.E., Williams, C.N., Wuertz, D.B., Vose, R.S. & et al. (2011) An overview of the Global Historical Climatology Network monthly mean temperature dataset, version 3. *Journal of Geophysical Research: Atmospheres*, 116, D19121. <https://doi.org/10.1029/2011JD016187>
- Lean, J. (2000) Evolution of the Sun's spectral irradiance since the Maunder Minimum. *Geophysical Research Letters*, 27, 2425–2428. <https://doi.org/10.1029/2000GL000043>
- Linán, I.D., Büntgen, U., González-Rouco, F., Zorita, E., Montávez, J., Gómez-Navarro, J. et al. (2012) Estimating 750 years of temperature variations and uncertainties in the Pyrenees by tree ring reconstructions and climate simulations. *Climate of the Past*, 8, 919–933. <https://doi.org/10.5194/cp-8-919-2012>
- Livezey, R.E., Masutani, M., Leetmaa, A., Rui, H., Ji, M. & Kumar, A. (1997) Teleconnective response of the Pacific-North American region atmosphere to large central equatorial Pacific SST anomalies. *Journal of Climate*, 10, 1787–1820. [https://doi.org/10.1175/1520-0442\(1997\)010<1787:TROTPN>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<1787:TROTPN>2.0.CO;2)
- Loader, N.J., Jalkanen, R., Mccarroll, D. & Moberg, A. (2011) Spring temperature variability in northern Fennoscandia AD 1693–2011. *Journal of Quaternary Science*, 26, 566–570. <https://doi.org/10.1002/jqs.1528>
- Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M. & Wanner, H. (2004) European seasonal and annual temperature variability, trends, and extremes since 1500. *Science*, 303, 1499–1503. <https://doi.org/10.1126/science.1093877>
- Manley, G. (1974) Monthly mean central England temperature, 1659–1973. *Quarterly Journal of the Royal Meteorological Society*, 100, 389–405. <https://doi.org/10.1002/qj.49710042511>
- Mann, M.E., Zhang, Z., Rutherford, S., Bradley, R.S., Hughes, M.K., Shindell, D. et al. (2009) Global signatures and dynamical origins of the Little Ice Age and Medieval Climate Anomaly. *Science*, 326, 1256–1260. <https://doi.org/10.1126/science.1177303>
- McCarroll, D., Loader, N.J., Jalkanen, R., Gagen, M.H., Grudd, H., Gunnarson, B.E. et al. (2013) A 1200-year multiproxy record of tree growth and summer temperature at the northern pine forest limit of Europe. *The Holocene*, 23, 471–484. <https://doi.org/10.1177/0959683612467483>
- Menne, M.J., Durre, I., Vose, R.S., Gleason, B.E. & Houston, T.G. (2012) An overview of the global historical climatology network-daily database. *Journal of Atmospheric and Oceanic Technology*, 29, 897–910. <https://doi.org/10.1175/JTECH-D-11-00103.1>
- Mikami, T. (1996) Long term variations of summer temperatures in Tokyo since 1721. *Geographical Reports of Tokyo Metropolitan University*, 31, 157–165.
- Moberg, A., Bergström, H., Kringsman, J.R. & Svanered, O. (2002) Daily air temperature and pressure series for Stockholm (1756–1998), in: *Improved Understanding of Past Climatic Variability from Early Daily European Instrumental Sources*, pp. 171–212, Dordrecht: Springer. https://doi.org/10.1007/978-94-010-0371-1_7
- Murphy, C., Broderick, C., Burt, T.P., Curley, M., Duffy, C., Hall, J. et al. (2018) A 305-year continuous monthly rainfall series for the island of Ireland (1711–2016). *Climate of the Past*, 14, 413–440. <https://doi.org/10.5194/cp-14-413-2018>
- Neter, J., Kutner, M.H., Nachtsheim, C.J. & Wasserman, W. (1996) *Applied linear statistical models*. 4, Irwin Chicago.
- Neukom, R., Gergis, J., Karoly, D.J., Wanner, H., Curran, M., Elbert, J. et al. (2014) Inter-hemispheric temperature variability over the past millennium. *Nature Climate Change*, 4, 362. <https://doi.org/10.1038/nclimate2174>
- Ohba, T., Morishima, W., Zaiki, M., Mikami, T. & Hirano, J. (2013) Reconstruction of July temperature variations since the 1830s in Kawanishi based on historical weather documents. *Japanese Progress in Climatology*, 87–90.
- Parker, D.E., Legg, T.P. & Folland, C.K. (1992) A new daily central England temperature series, 1772–1991. *International Journal of Climatology*, 12, 317–342. <https://doi.org/10.1002/joc.3370120402>
- Pauling, A., Luterbacher, J., Casty, C. & Wanner, H. (2006) Five hundred years of gridded high-resolution precipitation reconstructions over Europe and the connection to large-scale circulation. *Climate Dynamics*, 26, 387–405. <https://doi.org/10.1007/s00382-005-0090-8>
- Peterson, T.C. & Vose, R.S. (1997) An overview of the Global Historical Climatology Network temperature database. *Bulletin of the American Meteorological Society*, 78, 2837–2850. [https://doi.org/10.1175/1520-0477\(1997\)078<2837:A00TGH>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<2837:A00TGH>2.0.CO;2)
- Wetternachtersage. 500 Jahre Klimavariationen und Naturkatastrophen (1496–1995)*, Haupt-Verlag, 1999.
- Pongratz, J., Reick, C., Raddatz, T. & Claussen, M. (2008) A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochemical Cycles*, 22, GB3018. <https://doi.org/10.1029/2007GB003153>
- Popa, I. & Kern, Z. (2009) Long-term summer temperature reconstruction inferred from tree-ring records from the Eastern Carpathians. *Climate Dynamics*, 32, 1107–1117. <https://doi.org/10.1007/s00382-008-0439-x>
- Przybylak, R., Majorowicz, J., Wójcik, G., Zielski, A., Chorażyczewski, W., Marciniak, K. et al. (2005) Temperature changes in Poland from the 16th to the 20th centuries. *International Journal of Climatology*, 25, 773–791. <https://doi.org/10.1002/joc.1149>
- Rayner, N., Horton, E., Folland, C.K., Alexander, L.V., Rowell, D. et al. (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres*, 108, NO. D14, 4407. <https://doi.org/10.1029/2002JD002670>
- Rennie, J., Lawrimore, J., Gleason, B., Thorne, P., Morice, C., Menne, M. et al. (2014) The international surface temperature initiative

- global land surface databank: Monthly temperature data release description and methods. *Geoscience Data Journal*, 1, 75–102. <https://doi.org/10.1002/gdj3.8>
- Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M. et al. (2003) *The atmospheric general circulation model ECHAM 5. PART I: Model description*, Tech. rep. 349, Max Planck Institute for Meteorology.
- Rohde, R., Muller, R., Jacobsen, R., Muller, E., Perlmuter, S., Rosenfeld, A. et al. (2013) *A New Estimate of the Average Earth Surface Land Temperature Spanning 1753 to 2011*, *Geoinfor Geostat: An Overview*, 1, <https://doi.org/10.4172/2327-4581.1000101>
- Rohrer, M., Brönnimann, S., Martius, O., Raible, C.C., Wild, M. & Compo, G.P. (2018) Representation of extratropical cyclones, blocking anticyclones, and Alpine circulation types in multiple reanalyses and model simulations. *Journal of Climate*, 31, 3009–3031. <https://doi.org/10.1175/JCLI-D-17-0350.1>
- Rousseau, D. (2009) *Les températures mensuelles en région parisienne de 1676 à 2008*, La Météorologie, <https://doi.org/10.4267/2042/30038>
- Rousseau, D. (2013) *Les moyennes mensuelles de températures à Paris de 1658 à 1675: d'Ismail Boulliau à Louis Morin*, Rubrique: Climatologie, <https://doi.org/10.4267/2042/51098>
- Schmidt, N., Lipp, E., Rose, J. & Luther, M.E. (2001) ENSO influences on seasonal rainfall and river discharge in Florida. *Journal of Climate*, 14, 615–628. [https://doi.org/10.1175/1520-0442\(2001\)014<0615:EIOSRA>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<0615:EIOSRA>2.0.CO;2)
- Slonosky, V.C. (2002) Wet winters, dry summers? Three centuries of precipitation data from Paris. *Geophysical Research Letters*, 29, 34–41. <https://doi.org/10.1029/2001GL014302>
- Slonosky, V. (2014) Historical climate observations in Canada: 18th and 19th century daily temperature from the St. Lawrence Valley, Quebec. *Geoscience Data Journal*, 1, 103–120. <https://doi.org/10.1002/gdj3.11>
- Steiger, N.J., Smerdon, J.E., Cook, E.R. & Cook, B.I. (2018) A reconstruction of global hydroclimate and dynamical variables over the Common Era. *Scientific Data*, 5(1), 180086. <https://doi.org/10.1038/sdata.2018.86>
- Tardif, R., Hakim, G.J., Perkins, W.A., Horlick, K.A., Erb, M.P., Emile-Geay, J. et al. (2019) Last Millennium Reanalysis with an expanded proxy database and seasonal proxy modeling. *Climate of the Past*, 15, 1251–1273. <https://doi.org/10.5194/cp-15-1251-201>
- Townley, R. (1694) A letter from Richard Townley, of Townley in Lancashire, Esq; containing observations on the quantity of rain falling monthly, for several Years successively: communicated to the Royal Society. *Philosophical Transactions of the Royal Society of London*, 18, 51–58. <https://doi.org/10.1098/rstl.1694.0014>
- Townley, R. (1699) An account of what rain fell at Townly in Lancashire, in the years 1697, and 1698. with some observations on the weather; being part of a letter of the 12th of Jan. 1698. *Philosophical Transactions of the Royal Society of London*, 21, 47–48. <https://doi.org/10.1098/rstl.1699.0008>
- Valler, V., Brugnara, Y., Franke, J. & Brönnimann, S. (2020) Assimilating monthly precipitation data in a paleoclimate data assimilation framework. *Climate of the Past*, 16, 1309–1323. <https://doi.org/10.5194/cp-16-1309-2020>
- Valler, V., Franke, J. & Brönnimann, S. (2019) Impact of different estimations of the background-error covariance matrix on climate reconstructions based on data assimilation. *Climate of the Past*, 15, 1427–1441. <https://doi.org/10.5194/cp-15-1427-2019>
- Van Der Schrier, G. & Jones, P.D. (2008) Daily temperature and pressure series for Salem, Massachusetts (1786–1829). *Climatic Change*, 87, 499–515. <https://doi.org/10.1007/s10584-007-9292-x>
- Vinther, B.M., Andersen, K.K., Jones, P.D., Briffa, K.R. & Cappelen, J. (2006) Extending Greenland temperature records into the late eighteenth century. *Journal of Geophysical Research: Atmospheres*, 111, D11105. <https://doi.org/10.1029/2005JD006810>
- Wartenburger, R., Brönnimann, S. & Stickler, A. (2013) Observation errors in early historical upper-air observations. *Journal of Geophysical Research: Atmospheres*, 118, 12–012. <https://doi.org/10.1002/2013JD020156>
- Weselowskij, K. (1857) *O Klimatje Rassii, Imperatorskaya Akademiya Nauk*, St. Petersburg.
- Wheeler, D.A. (2006) The Gibraltar climatic record: Part 1-the history of weather observations. *Weather*, 61, 36–39. <https://doi.org/10.1256/wea.74.05>
- Wheeler, D. (2007) The Gibraltar climatic record: Part 2-precipitation. *Weather*, 62, 99–104. <https://doi.org/10.1002/wea.49>
- Whitaker, J.S. & Hamill, T.M. (2002) Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, 130, 1913–1924. [https://doi.org/10.1175/1520-0493\(2002\)130<1913:EDAWP>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWP>2.0.CO;2)
- Wilks, D.S. (2011) *Statistical Methods in the Atmospheric Sciences*. 100, Academic press.
- Wilson, R., Anchukaitis, K., Briffa, K.R., Büntgen, U., Cook, E.R., D'Arrigo, R. et al. (2016) Last millennium northern hemisphere summer temperatures from tree rings: Part I: The long term context. *Quaternary Science Reviews*, 134, 1–18. <https://doi.org/10.1016/j.quascirev.2015.12.005>
- WSL. (2019) *MXD chronologies*, Swiss Federal Institute for Forest, Snow and Landscape Research, <https://www1.ncdc.noaa.gov/pub/data/paleo/treering/updates/schweingruber/> accessed 8 Dec 2019.
- Yoshimori, M., Raible, C.C., Stocker, T.F. & Renold, M. (2010) Simulated decadal oscillations of the Atlantic meridional overturning circulation in a cold climate state. *Climate Dynamics*, 34, 101–121. <https://doi.org/10.1007/s00382-009-0540-9>
- Zaiki, M., Können, G.P., Tsukahara, T., Jones, P.D., Mikami, T. & Matsumoto, K. (2006) Recovery of nineteenth-century Tokyo/Osaka meteorological data in Japan. *International Journal of Climatology*, 26, 399–423. <https://doi.org/10.1002/joc.1253>
- Zambri, B., Robock, A., Mills, M.J. & Schmidt, A. (2019) Modeling the 1783–1784 Laki eruption in Iceland: 2. *Climate impacts*, *Journal of Geophysical Research: Atmospheres*, 124, 6770–6790. <https://doi.org/10.1029/2018JD029554>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Valler V, Franke J, Brugnara Y, Brönnimann S. An updated global atmospheric paleo-reanalysis covering the last 400 years. *Geosci Data J.* 2021;00:1–19. <https://doi.org/10.1002/gdj3.121>